7-2-2020

Patterns of Population Displacement During Mega-Fires in California detected using Facebook Disaster Maps

Shenyue Jia

Seung Hee Kim

Son V. Nghiem

Paul Doherty

Menas Kafatos

Patterns of Population Displacement During Mega-Fires in California detected using Facebook Disaster Maps

Patterns of population displacement during mega-fires in California detected using Facebook Disaster Maps

IOP Publishing Environ. Res. Lett. 15 (2020) 074029 https://doi.org/10.1088/1748-9326/ab8847

Environmental Research Letters

Patterns of population displacement during mega-fires in California detected using Facebook Disaster Maps

Shenyue Jia[1], Seung Hee Kim[1,4], Son V Nghiem[2], Paul Doherty[3] and Menas C Kafatos[1]

[1] Center of Excellence in Earth Systems Modeling and Observations (CEESMO), Chapman University, Orange, California, United States of America
[2] NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, United States of America
[3] National Alliance for Public Safety GIS (NAPSG) Foundation, Washington D.C., United States of America
[4] Author to whom any correspondence should be addressed

E-mail: sekim@chapman.edu

Keywords: Facebook disaster maps, crowdsourced data, social media, Mann-Kendall trend, anomaly analysis, wildfires, California

OPEN ACCESS

RECEIVED
30 January 2020

REVISED
30 March 2020

ACCEPTED FOR PUBLICATION
9 April 2020

PUBLISHED
2 July 2020

Abstract

The Facebook Disaster Maps (FBDM) work presented here is the first time this platform has been used to provide analysis-ready population change products derived from crowdsourced data targeting disaster relief practices. We evaluate the representativeness of FBDM data using the Mann-Kendall test and emerging hot and cold spots in an anomaly analysis to reveal the trend, magnitude, and agglommeration of population displacement during the Mendocino Complex and Woolsey fires in California, USA. Our results show that the distribution of FBDM pre-crisis users fits well with the total population from different sources. Due to usage habits, the elder population is underrepresented in FBDM data. During the two mega-fires in California, FBDM data effectively captured the temporal change of population arising from the placing and lifting of evacuation orders. Coupled with monotonic trends, the fall and rise of cold and hot spots of population revealed the areas with the greatest population drop and potential places to house the displaced residents. A comparison between the Mendocino Complex and Woolsey fires indicates that a densely populated region can be evacuated faster than a scarcely populated one, possibly due to better access to transportation. In sparsely populated fire-prone areas, resources should be prioritized to move people to shelters as the displaced residents do not have many alternative options, while their counterparts in densely populated areas can utilize their social connections to seek temporary stay at nearby locations during an evacuation. Integrated with an assessment on underrepresented communities, FBDM data and the derivatives can provide much needed information of near real-time population displacement for crisis response and disaster relief. As applications and data generation mature, FBDM will harness crowdsourced data and aid first responder decision-making.

1. Introduction

Abrupt population displacement is a common emergency response strategy applied before or during a natural disaster. Evaluating the effectiveness of this strategy and obtaining useful information for disaster relief during a large population displacement is of great interest to both policy-makers and researchers. The accessibility of cellular tower transmission data with geospatial signature has been greatly constrained amid the concern of data abuse and privacy preservation [1, 2]. Therefore, many researchers and practitioners have turned to geolocation data that are crowdsourced through leading social media platforms to assimilate population movements in a timely manner and assist disaster relief. Crowdsourced geolocation data through location-based services are sensitive to the fluctuation of users' position change and respond promptly to the spatial movement and formation of clusters. For the most popular social media platforms like Twitter and Facebook, their crowdsourced geolocation data can serve as proxies of population distribution to answer the most basic yet important question for any crisis response task: 'Where are the people?'

Studies and actual practices have been carried out using the geotagged Tweets and Flickr picture uploads to track and reveal the change and distribution of population during natural disasters and public safety crises [3–8]. These successful applications

also motivated the leading social media platforms to develop their official crisis response tools. Such efforts have greatly increased the popularity of crowdsourced data in time-sensitive situations for urgent emergency actions, with natural disasters being prime examples. While the representativeness of the social media population has been questioned [4, 9, 10], the steadily increasing pool of active users of these platforms [11] has helped capture a more accurate picture of human activities [12].

Current limitations to advance the use of crowdsourced data in emergency response can be framed into three major points. First, most studies using crowdsource data from social media, especially Twitter [5], were focused on the development of algorithms and pipelines to effectively retrieve and clean the raw data from platforms but could not derive meaningful metrics to support decision-making [13–15]. Second, pipelines of data processing applied by different teams sometimes varied, causing difficulties to compare the derived outcomes. Third, a knowledge detachment occurred between data analysts and first responders due to the different backgrounds and focus of interest [16].

These limitations can partly be alleviated by developing a universal and analysis-ready dataset of population movements for major disasters and emergencies across different parts of the world. Such a dataset can relieve analysts the complex preprocessing of crowdsourced data and support immediate analyses to decision-making as well as it can be compared across different situations. The launch of Disaster Maps by Facebook in 2017 is a major effort to address this need. As a part of Facebook Data for Good (dataforgood.fb.com) initiative, Facebook Disaster Maps (FBDM) tool provides population displacement information every 8 h in a 1 km grid during any major natural disaster across the globe [17]. Anonymized and aggregated based on the crowdsourced data from Facebook App and Facebook Safety Check App usage with location service enabled, FBDM not only provides population count changes over the period of natural disasters, but also generates metrics to capture population anomaly compared with a pre-crisis situation. In addition, FBDM offers network and power coverage maps to illustrate the Internet and power connectivity of users, based on the anonymized connection data through cellular network to Facebook server and the information of device charging from users [17]. For the first time, an analysis-ready and near real-time population change product has become available to assist disaster relief works, including identifying the places of population aggregation during a disaster [18], retrieving the directions of population displacement [19], and evaluating the degree of resource needs during a disaster [20]. Leveraging this dataset with precaution considerations of privacy preservation not only assists disaster relief works during particular disasters but also helps to identify disaster-prone communities and plans for effective mitigation strategies.

In this paper, we use the FBDM population change data to analyze the spatio-temporal patterns of population displacement during the Mendocino Complex and the Woolsey fires, two mega-fires which occurred in California, USA, in 2018. Our study first evaluates the representativeness of FBDM population, then calculates a series of derivatives from the basic FBDM population change data to reveal the spatio-temporal pattern of population change as well as hot spots. Furthermore, this study assesses the versatility of these derivatives by comparing the outcomes between mega-fires occurred in scarcely and densely populated regions.

2. Mendocino Complex fire and Woolsey fire

The Mendocino Complex (459 000 acres) and Woolsey fires (96 949 acres) were selected due to their exceptionally large perimeter and damage in California during the 2018 fire season, the first full-length fire season after the launch of FBDM. These two fires also represented the environmental and social settings of fire-prone areas in less populated, forest covered Northern California versus heavily populated, shrubland dominated Southern California.

As the largest wildfire recorded in California's history, the Mendocino Complex fire started on 27 July 2018 and was not fully contained until 7 November 2018. This fire complex consists of two fires, the smaller, earlier contained River fire and the larger Ranch fire (figure 1(a)). Mandatory evacuations initiated and ended on different days near the perimeters of the River and Ranch fires (figure 2(a)). For the River fire, mandatory evacuation orders were placed shortly after its ignition on July 27 and lifted on July 31, which was before the situation escalated and the release of FBDM dataset. For the Ranch fire, mandatory orders were in place from August 3 to August 7 for the neighborhoods on the west side of the downwind direction. The burned area has a low population density at 76 people per km^2, thus causing less impacts than fires with similar size (table 1).

In contrast, the Woolsey fire broke out in a heavily populated wildland urban interface (WUI) close to major metropolitan areas of Los Angeles and Ventura County in Southern California (figure 1(b)), and destroyed many more properties than the Mendocino Complex fire (table 1). The fire quickly expanded southbound towards the Pacific coast by the strong Santa Ana winds over the shrubland. Mandatory evacuation orders were placed on 8 November 2018 for inland and coastal cities but were lifted at different times (figure 2(b)). The inland cities near the origin of fire lifted the mandatory evacuation orders on November 10, while cities located in the downward wind direction and the outskirts of fire perimeter kept

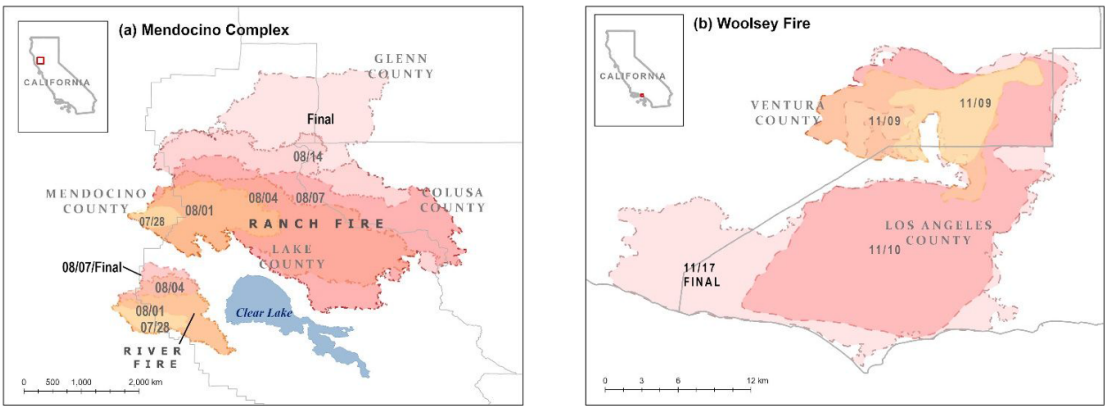

Figure 1. Progression and final perimeters of the Mendocino Complex fire (a) and the Woolsey fire (b). The change of fire perimeters as the burning continued is illustrated using boundaries with different colors and labels of dates. Note that maps are shown in different scales.

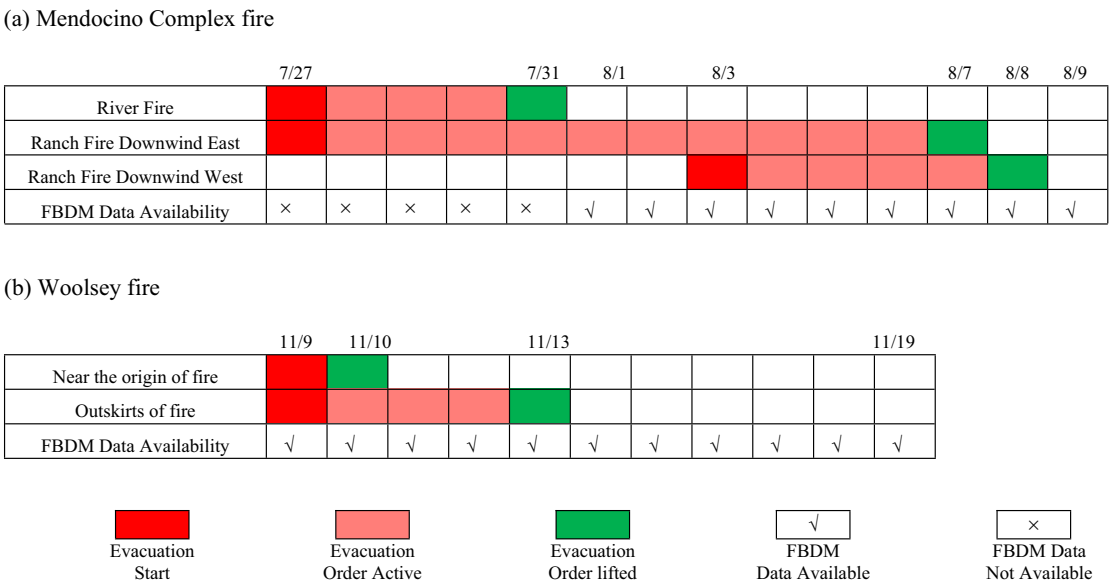

Figure 2. Timeline of fire progression and the placing/lifting of evacuation orders in and near the perimeters for Mendocino Complex fire (a) and Woolsey fire (b). Facebook Disaster Maps (FBDM) data availability is shown at the bottom of panel a and b.

Table 1. Basic information of Mendocino Complex fire and Woolsey fire (citation of data).

Fire name	Dates	Burned area	Land cover	Property destroyed	Fatalities	Population Density
Mendocino Complex	Jul. 27-7 November 2018	1858 km^2	Temperate and boreal forest	280	1	16 per km^2
Woolsey	Nov. 8-21 November 2018	392 km^2	Shrubland	1643	3	335 per km^2

the evacuation orders active until November 13 or later.

3. Data and method

3.1. Data

In this study, we focus on two key FBDM metrics, the Facebook baseline population (FBP) for pre-crisis situation and Facebook population z-score for anomalies during crisis. FBP is calculated based on the multi-day average login and usage of Facebook at a geographical location and a given time of the day over 5–13 weeks of the pre-crisis situation [17]. Z-score is a standardized and unit-less metric derived to measure the departure of crisis population from the pre-crisis situation [18]. Signs of z-score indicate the departure below (negative) or above (positive) the pre-crisis level. Z-score is comparable across the geographical location and time. A greater absolute value of z-score indicates a larger departure from the pre-crisis level. Both FBDM metrics are available every 8 h during the major disasters and posted in

1-km grids to aggregate and anonymize the individual users [18]. More information about FBDM data generation and accessibility is included in supplemental information section 2.

In order to assess the representativeness of FBDM data, we employed two products of total population, the 2018 American Community Survey (ACS) population estimates at the zip code level from the U.S. Census as well as the Gridded Population of the World (GPW) version 4 (1-km grid) of the Columbia University [21].

We obtained the progression of fires and final fire perimeters from the USGS Geospatial Multi-Agency Coordination (GeoMAC) and the California Department of Forestry & Fire (CAL FIRE). Information of local government announcements on evacuation orders was from CAL FIRE and InciWeb (https://inciweb.nwcg.gov/). Geographical boundaries from U.S. Census TIGER were used to assist the analysis.

3.2. Calculating the penetration rate of Facebook users

To assess the reliability of results drawn from FBDM data, we first evaluated the data representativeness using the penetration rate of Facebook service into the total population. Penetration rate is defined as below (equation (1)).

$$Penetration\,Rate = \frac{Number\,of\,regular\,Facebook\,users}{Total\,population} \times 100\% \quad (1)$$

In this study, we calculate two sets of penetration rates using different total population data, GPW in a 1-km grid and ACS population estimates at the zip-code level. For the penetration rate based on GPW, we resample the GPW data to the grid footprint of FBP to ensure the two datasets were co-registered. For the penetration rate based on ACS estimates, we identify FBDM grids inside each zip code and sum total the FBP values as the number of regular Facebook users. Penetration rates were calculated for all 36 time slices from the FBDM Woolsey fire dataset. We also calculate the zip code-level average for GPW-based penetration rate to better compare it with the ACS-based counterpart. We average the penetration rate of all time slices and every time stamp to reveal the overall penetration level of Facebook usage and its diurnal cycle.

3.3. Deriving trends and emerging hotspots for Facebook population z-score

We incorporated the space-time cube [22], a data structure designed to store spatial and temporal information to facilitate the analysis of population displacement pattern in space and time, including temporal trend, clustering of high and low values, and the evolution of the clusters. We construct the space-time cube by laying over the z-score map of every time stamp from the earliest to the latest.

Wildfire emergencies usually involve multiple critical events like the initiation and termination of mandatory evacuation orders. To reveal the change of population toward each of these events, we separate the z-score time series into sections based on the timing of these events and calculate the trend of z-score time series for each section using Mann-Kendall test, one of the widely used distribution-free tests of trend in time series analysis to reveal monotonic trend [23–25]. We also examine the temporal change of z-score over the entire space-time cube and detect the tipping points when the direction of trend changed. We then compare the detected tipping points with the official timing of the placing and lifting of evacuation orders. Such analysis is conducted for the z-score time series at each 1-km hexagon in the fire affected area to reveal detailed spatial similarity or difference of population displacement, which can only be made possible with the crowdsourced data. If population change during the crisis presents as clusters in space, we calculate average z-score over each cluster to identify the intra-cluster similarity and inter-cluster distinction.

We conducted emerging hot spot analysis to detect trends in the clustering of z-score. The clustering is represented as hot and cold spots, or the aggregation of high and low population grids in space. During wildfires, such clustering in space might grow or shrink, indicating the dynamic of population displacement in space and time. To account for this process, we calculated the Getis-Ord Gi* statistics [26, 27] to detect hot and cold spots over the 1-km resolution z-score map at each time slice. Then we conduct the Mann-Kendall trend test for the Gi* statistics of all time slices at each 1-km grid. The Mann-Kendall trend of Gi* is calculated to categorize each 1-km grid as a new, unstable, or stable hot and cold spot, or no significant spatial aggregation. Such analysis replicates the Emerging Hot Spot Analysis from the Space Time Pattern Mining Toolbox of ArcGIS [28].

4. Results

4.1. Representativeness of Facebook disaster maps population data

The penetration rate of Facebook usage varied by region. Regardless of the total population data source used in calculation, the penetration rate was higher in densely populated areas, especially along the major transportation corridors (figure 3). In less populated and remote areas near mountains and forests, the penetration rate of Facebook users was much lower. However, some spuriously high penetration grids also are present in remote

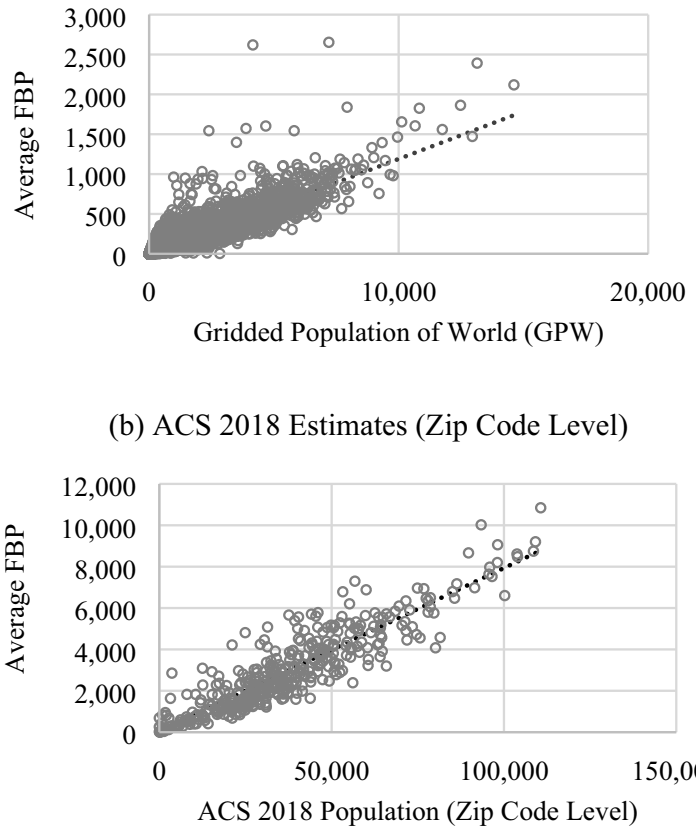

Figure 3. Spatial distribution of Facebook penetration rate based on the Facebook Disaster Maps (FBDM) Woolsey fire dataset. Panel a represents the penetration rate calculated based on GPW posted to 1 km grids. Panel b shows the zip code level average penetration rate based on GPW. Panel c shows the penetration rate calculated based on zip code level ACS 2018 population.

areas (figure 3(a)) due to the noise in the original FBP data. Such an uncertainty needs to be considered while interpreting results derived for studies about wildfires, which often affect remote and less populated areas. In addition, areas with a higher proportion of younger population shows a higher penetration rate of Facebook users, indicating a sampling bias of the FBDM data (SI figure 1 https://stacks.iop.org/ERL/15/074029/mmedia). The penetration rate of Facebook usage also indicates a diurnal cycle, which plunged after midnight and peaked in the mid-morning (SI figure 2).

FBP shows a high correlation with the total population. In the FBDM Woosley fire dataset, the R^2 between the average FBP and the total population from GPW and ACS 2018 estimates was 0.8322 and 0.8171, respectively (figure 4). The average penetration rate, shown as the slope of the linear fit equation in figure 4, indicated that FBP sampled 11.89% from the GPW and 7.91% from the ACS 2018 population estimates in the FBDM Woolsey fire dataset. GPW-based penetration rate was higher than the ACS-based rate, due to a systematical underestimation of population in GPW [29].

4.2. Overall and regional population dynamics

We found that the overall and regional average z-score captured the tipping points of population when relevant communities responded to the placing and lifting of mandatory evacuation orders in both cases. The dynamics of z-score involved three stages: 1) drastic decrease after the evacuation order, with some significant increase near open shelters or nearby towns, 2) prominent increase after the lifting of evacuation order, and 3) slowing increase until z-score reached zero. The speed and magnitude of changes during each stage varied by location. During the Mendocino Complex fire, z-score took 9 time steps (i.e. 72 h, time step 3 to 11) to decrease to its lowest since the placing of evacuation order (figure 5, blue line). For the Woolsey fire, such time was much shorter (32 h, time step 5 to 9) (figure 5, orange line). The lowest z-score might stay for several time steps, depending on how soon the evacuation orders were lifted. Upon the lifting of evacuation orders, z-score climbed up in a slower pace than its decrease to reach zero (figure 5), representing overall situation of population distribution.

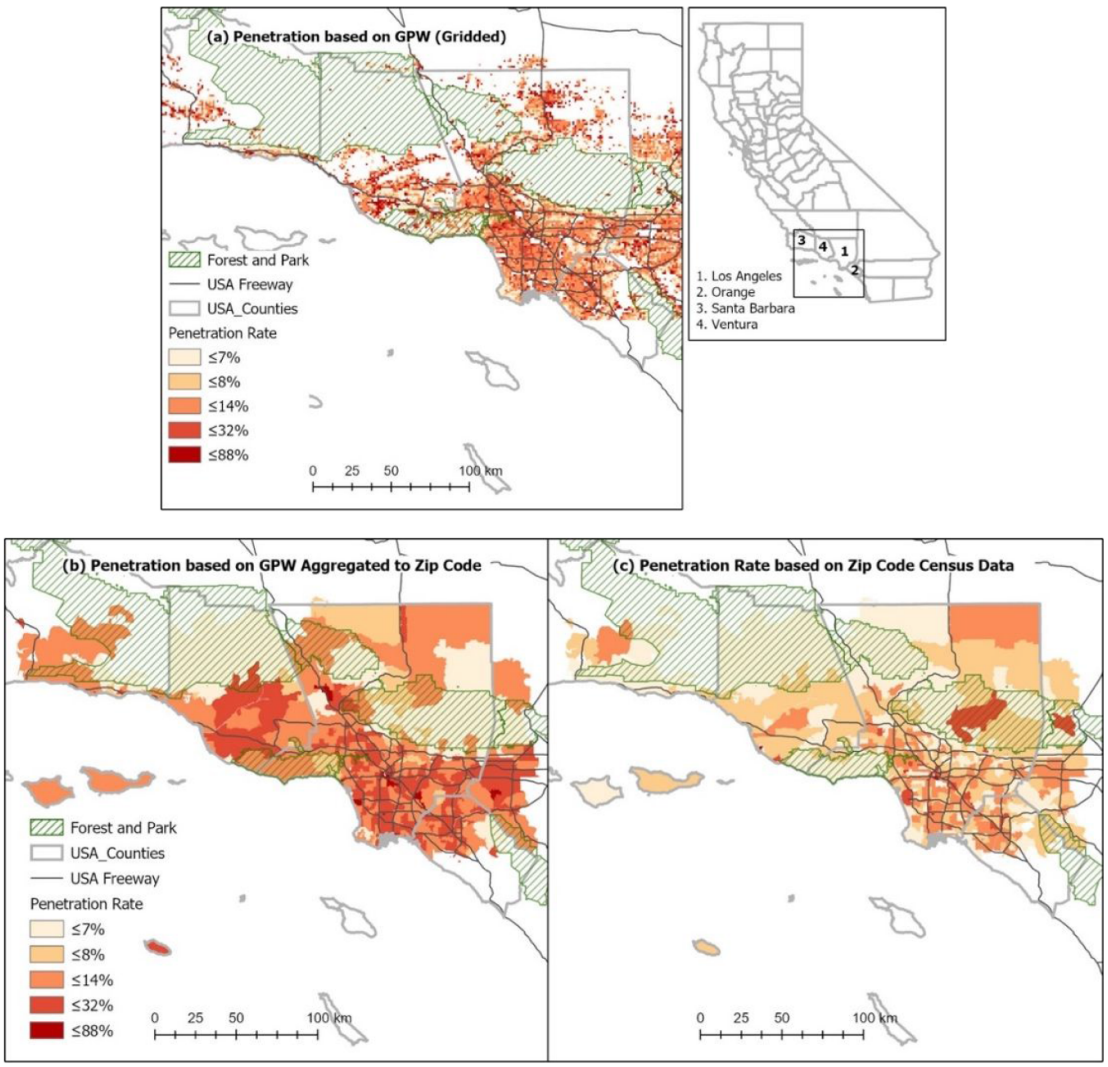

Figure 4. Average Facebook baseline population (FBP) from the FBDM Woolsey fire dataset and the total population. GPW 1 km population and ACS 2018 zip code level estimates were used in panel a and b as the total population.

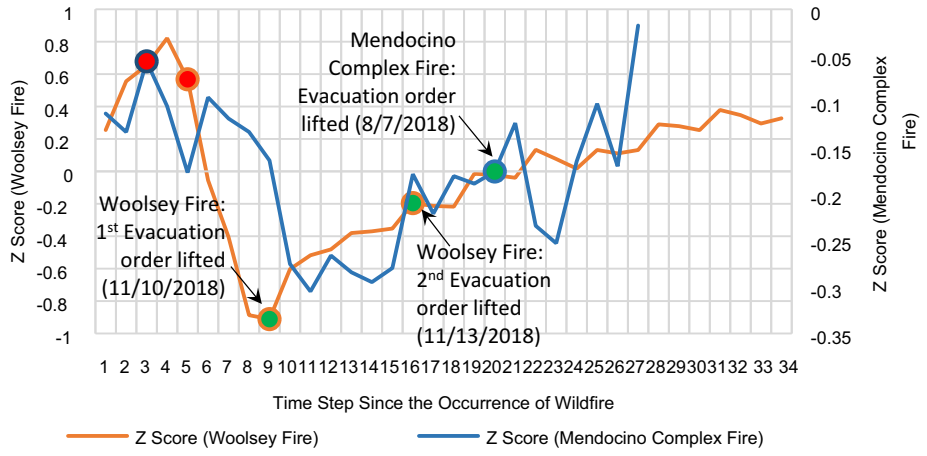

Figure 5. Facebook Disaster Maps (FBDM) z-score dynamics inside the entire region of interest affected by fire during the fire progression. Blue and orange lines represent the Mendocino Complex fire and the Woolsey fire, respectively. Red circles indicate the issue of an evacuation order. Green circles indicate evacuation order lifting. Z-scores are presented in an 8-hour time step. Positive z-score indicates that Facebook users are above the pre-crisis average, while negative z-score indicates that Facebook users are below the pre-crisis average.

The average z-score also captured the spatial difference emerging from the differing time of evacuation orders placing and lifting by places. During the Woolsey fire, evacuation orders were lifted much sooner for neighborhoods near the fire origin than elsewhere. Such a difference resulted in three prominent types of population dynamic patterns in the fire-affected region, including area of quick return, slow return, and flash increase (figure 6). A subtle yet distinguishable difference within each type also indicates the linkage between the size of area affected and the magnitude of population movement. With a much smaller area and fraction of residents affected by the mandatory evacuation, Thousand Oaks showed a less drastic change of z-score than other neighborhoods in the quick return group, where the majority of the neighborhood was evacuated (figure 6(b)). A greater area exposed to wildfire progression and evacuation meant greater damage and longer road closure. Such inconveniences resulted in the low possibility and willingness to return, therefore led to a more gradual z-score recovery (Calabasas and Hidden Hills) or even no recovery at all (Malibu) (figure 6(c)). Aside from the decline of population, z-score also pinpointed possible destinations of the displaced people. Z-score increased slightly in November 9 outside the northern boundary of fire perimeters and quickly declined after November 10 (figure 6(d)), in parallel to the placing and lifting of the evacuation orders in Thousand Oaks and nearby neighborhoods (figure 6(b)). As the closest populated area not affected by the Woolsey fire in this region, these neighborhoods became potential destinations to accommodate the displaced people from Thousand Oaks and nearby places.

4.3. Grid-level trends in different stages of wildfire emergencies

Trends of z-score at each grid showed a consistent pattern with the overall and regional average values. During the Woolsey fire, trends showed a more homogeneous pattern in space as the evacuation initiated concurrently across the region and terminated simultaneously within nearby neighborhoods. Most areas affected by the Woolsey fire showed a significant decrease of z-score while people were fleeing the evacuated area, with scattered increasing grids near some open shelters (figure 7(a)). When the lifting of evacuation order started, the spatial aggregation of increased z-score was prominent (figure 7(b)). Such increase expanded as the lifting continued (figure 7(c)). In contrast, open shelters showed a significant decrease as the lifting of evacuation continued (figures 7(b) and (c)). For the Mendocino Complex fire, the trend of z-scores also captured the population changes in areas undergoing different types of population displacement. As evacuation orders to the west of the River fire perimeter were lifted before the FBDM data became available, this area showed significant increase of z-score while areas to the south of the Ranch fire was undergoing a population decrease after the newly placed evacuation order (figure 8(a)). When the evacuation orders near Ranch fire were lifted, positive trend occurred in parallel to the z-score decrease near the two OPEN shelters in the southeast corner of Clear Lake (figure 8(b)). Overall, the trend straightforwardly showed the most active areas of population decrease and increase during wildfire emergencies, providing insights on population displacement.

4.4. Changing hot and cold spots of population displacement

As spatial clustering of population anomaly, hot and cold spots of z-score might expand, persist, shrink, or disappear as a wildfire emergency unfolds. Recording the rise and fall of them not only documents the degree of population displacement but also identifies the temporal places of the population. During an active evacuation period, cold spots were consecutive and persistent across the evacuated area, with new cold spots popping up near its outskirt (figures 9(a) and 10(a)). When the evacuation orders were lifted, cold spots either became less stable or disappeared (figures 9(b) and 10(b)), which was consistent with the increasing trend of z-score (figures 7(b) and 8(b)). When most residents returned home, a large number of cold or hot spots became unstable or even disappeared (figures 7(c) and 8(b)), indicating the end of the anomalous population change arising from the displacement. Compared to cold spots, hot spots were less common while the population displacement was active. Yet, they represented possible places where displaced people were housed, providing crucial information for disaster relief works. A series of stable hot spots arose at or near the open shelters and towns not affected by the fire while the evacuation orders were in place (figures 9(a), (b) and 10(a)). New and unstable hot spots appeared in areas experiencing a strong wave of population return after evacuation ended (figures 9(c) and 10(a) near the River fire). Meanwhile, some consecutive and persistent hot spots near open shelters disappeared as people housed in these facilities and neighborhoods left for home (figure 10(b)). The rise and fall of hot spots near open shelters were more prominent during the Mendocino Complex fire, partly due to the fewer options people have in this less populated, relatively isolated region in Northern California.

5. Discussion and conclusions

FBDM data could reveal the trends, magnitude, and spatial clustering of population displacement in a timely manner with adequate representativeness out of the total population in area with intensive usage such as California. Though biased in age and other demographic factors, FBDM promptly captured the

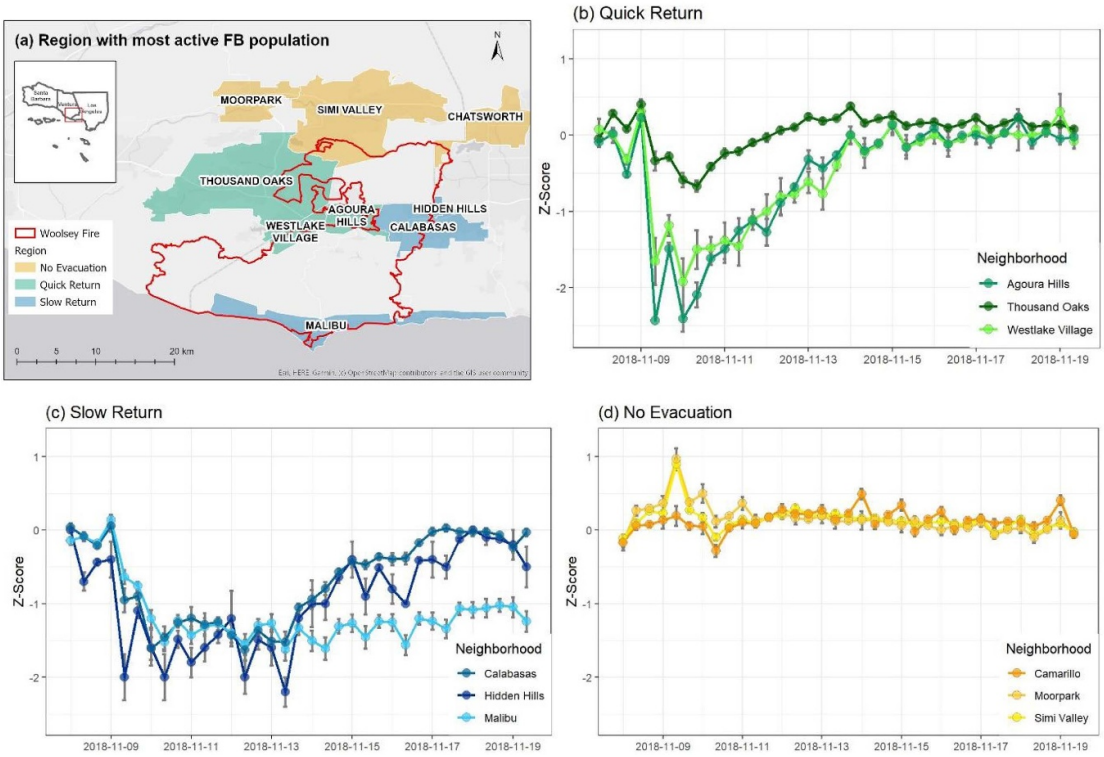

Figure 6. Facebook Disaster Maps (FBDM) z-score change of every 8 h during the progression of Woolsey Fire for different locations due to the different fire progression extent. Green, blue, yellow indicated neighbourhoods with quick return, slow return, and no evacuation. Line charts from panel b to d represent the temporal change of z-score the three types of neighbourhoods mentioned above. Error bars are plotted using ±standard error. Positive z-score indicates that Facebook users are above pre-crisis level, while negative z-score indicates that Facebook users are below pre-crisis level.

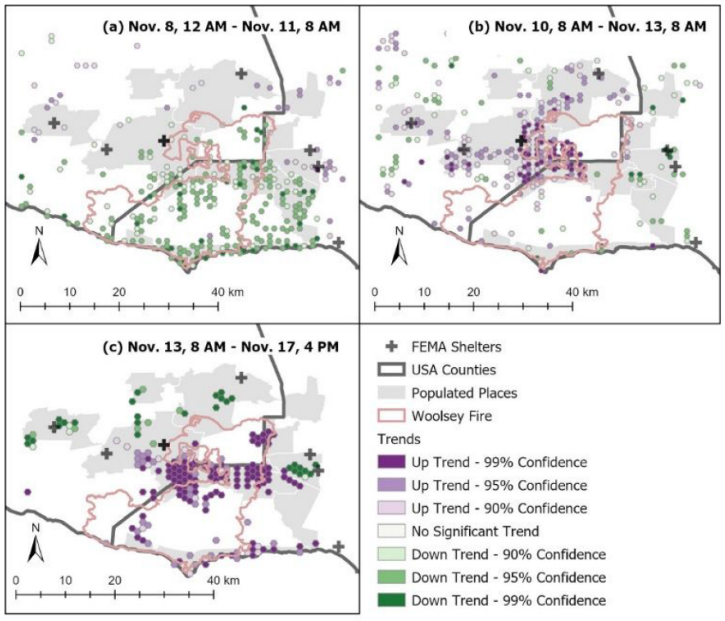

Figure 7. Significant trend of z-score in and adjacent to the Woolsey fire perimeter when communities responded to evacuation placing and lifting. Panels illustrate three stages during the fire emergency: (a) during the significant population drop after the placing of evacuation order, (b) while the evacuation orders were active or partly lifted, and (c) after evacuation orders were completely lifted (c).

abrupt population displacement when the placing and lifting of mandatory evacuation orders occurred. The dataset was sensitive not only to the prominent decrease of population in the evacuated area but also to the significant and flashy increase of population resulted from the displacement. The timely response from the platform and the provision of comparable and analysis-ready products provides a foundation for trend analysis and spatial pattern detection to support decision making. State and local

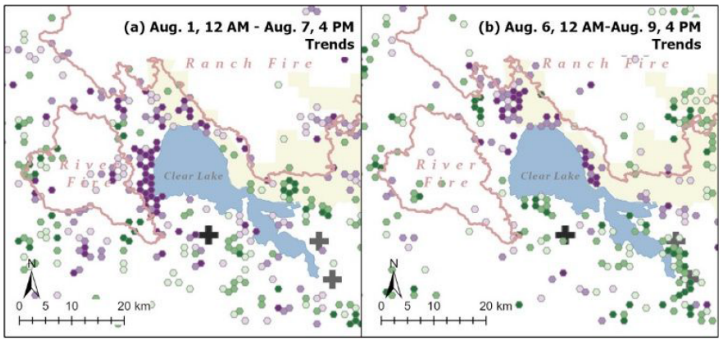

Figure 8. Significant trend of z-score in and adjacent to the Mendocino Complex fire perimeter when communities responded to evacuation placing and lifting. Panels illustrate two stages during the fire emergency: (a) while the evacuation orders were active or partly lifted and (b) after evacuation orders were completely lifted. Legend is the same as figure 7.

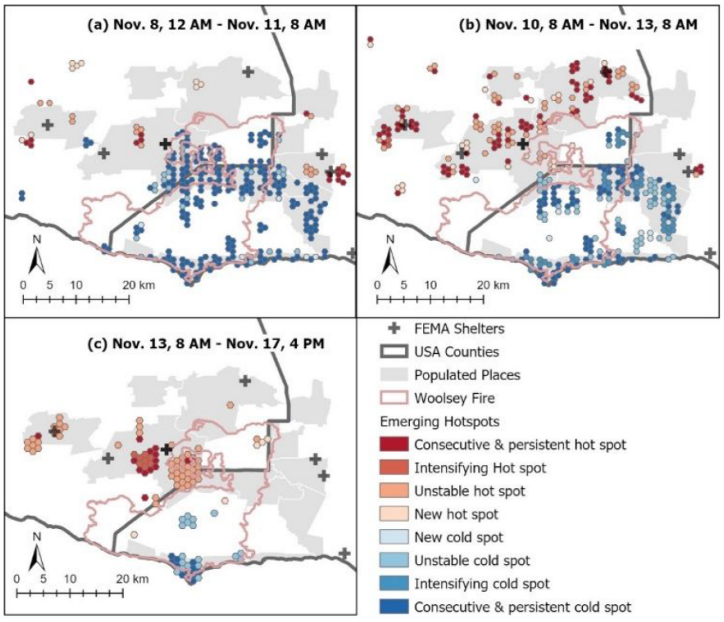

Figure 9. Emerging hot and cold spots of the Woolsey fire when communities responded to evacuation placing and lifting. Panels illustrate three stages during the fire emergency: (a) during the significant population drop after the announcement of evacuation order, (b) while the evacuation orders were active or partly lifted, and (c) after evacuation orders were completely lifted.

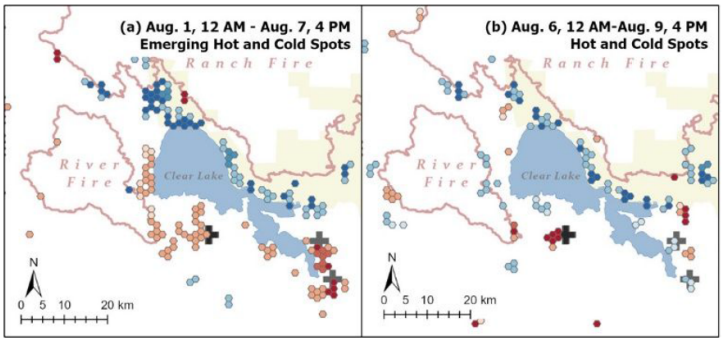

Figure 10. Emerging hot and cold spots of the Mendocino Complex fire when communities responded to evacuation placing and lifting. Panels illustrate two stages during the fire emergency: (a) while the evacuation orders were active or partly lifted and (b) after evacuation orders were completely lifted. Legend is the same as figure 9.

agencies design information products that could be used for monitoring local events like wildfires. We believe FBDM can improve usability of the platform by increasing the frequency at which data are released and by partnering with organizations who can produce actionable information products.

Distilling map series of population into meaningful secondary derivatives was necessary to harness the power of FBDM data. In this study, we calculated the Mann-Kendall trend and derived emerging hot and cold spots of Facebook population z-score to reveal the spatio-temporal pattern of population displacement. The rise of persistent and intensifying hot spots near open shelters during the evacuation was crucial information for disaster relief practice and post-disaster strategy. The analytical pipeline is highly transferrable to other disasters or emergencies with a better predictability, such as tornadoes, hurricanes and planned power outage amid extreme wildfire risk exacerbated by urbanization.

Comparability, a crucial feature of FBDM data, is critical for post-disaster reflection and future planning improvements. The same protocol of data generation enables the intercomparison between fires occurring in regions with varying social and environmental settings. In this study, we found that the evacuation in densely populated fire-prone area was faster due to the size of geographical area involved and the connectivity to transportation network. Population density affects the needs of resources and a less populated region might need resources prioritized to open shelters. In contrast, open shelters were not identified as hot spots in densely populated regions, as the displaced residents may have alternative options for temporary stay by exploiting their social connections [30] in nearby communities.

FBDM under-samples the elder population due to the relatively low popularity of Facebook in these age groups. Although such bias does not significantly affect the detection of the overall displacement trend and pattern, it could overlook the communities demanding more assistance yet failing to voice themselves out in the FBDM data. An analysis of social vulnerability based on the latest demographic data would be necessary [31] to identify the underrepresented but highly vulnerable communities. Such a discrepancies should be taken into consideration while using FBDM data to evaluate whether a neighborhood has been sufficiently evacuated and advising first responders to target the most vulnerable neighborhoods. In addition, the Internet connectivity can be disrupted during emergency situations, leading to an underestimation of human activities [32]. The connectivity of mobile network map from FBDM [17] should be employed to account for such uncertainty and identify places with a higher possibility for a false signal of population patterns.

Challenging the so-called 'The Tyranny of the Tweet' [33] in a crisis response sector, FBDM is the first platform to provide analysis-ready products from crowdsourced data for disaster relief in a timely manner. We have found that FBDM population z-score is a useful metric to reveal the temporal trend and spatial clustering during the massive population displacement arising from major disasters such as mega-fires. Derivatives from z-score, such as the Mann-Kendall trend, hot/cold spots, and the change of hot/cold spots can provide insightful information to evaluate the progress of population evacuation and return, identify agglomeration of displaced population, and assess the effectiveness of disaster relief strategy. With an uncertainty evaluation on the underrepresented groups, an improved workflow, and a coupled analysis with auxiliary products (movement vectors, mobile connectivity), FBDM population can help first responders deliver assistance to much needed communities and assist decision-makers to mitigate the disorder and chaos during crisis response. As illustrated in this paper, the FBDM method is versatile and thus adaptable to extreme wildfire situations such as the recent bushfires in Australia that have caused record-breaking population displacement [34]. In short, future research should focus on providing FBDM in 'the right format, to the right people, at the right time' in order to allow decision-makers and first responders to enhance outcomes for survivors.

Regarding future research, together with FBDM or similar crowdsourced data and independent remote sensing observations, the method presented in this paper can be adapted for applications to biological disasters in assessing the efficacy of emergency measures, such as the quarantine or lockdown orders amid the COVID-19 pandemic [35].

Acknowledgments

We thank Alex Pompe and Lauren McGorman from Facebook Data for Good for granting the access to Facebook Disaster Maps to conduct this study. We also thank Andrew Schroeder from Direct Relief for his input on conceptualizing the research ideas. The research, including urbanization impacts and population dynamics, carried out at the Jet Propulsion Laboratory, California Institute of Technology, was supported by the National Aeronautics and Space Administration (NASA) Land-Cover and Land-Use Change (LCLUC) Program.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analysed in this study.

References

[1] Deville P, Linard C, Martin S, Gilbert M, Stevens F R, Gaughan A E *et al* 2014 Dynamic population mapping using

mobile phone data *Proc. Natl. Acad. Sci.* **111** 15888–93
[2] Blondel V D, Decuyper A, and Krings G 2015 A survey of results on mobile phone datasets analysis *Epj Data Sci.* **4** 1
[3] Reuter C, Hughes A L, and Kaufhold M A 2018 Social media in crisis management: an evaluation and analysis of crisis informatics research *Int. J. Hum. Comput. Int.* **34** 280–94
[4] Zou L, Lam N S, Shams S, Cai H, Meyer M A, Yang S *et al* 2019 Social and geographical disparities in twitter use during hurricane harvey *Int. J. Digit. Earth* **12** 1300–18
[5] Martín Y, Li Z, and Cutter S L 2017 Leveraging twitter to gauge evacuation compliance: spatiotemporal analysis of hurricane Matthew *PLoS One* **12** e0181701
[6] Cai H, Lam -N S-N, Zou L, Qiang Y, and Li K 2016 Assessing community resilience to coastal hazards in the lower mississippi river basin *Water* **8** 46
[7] Guan X, and Chen C 2014 Using social media data to understand and assess disasters *Nat. Hazards* **74** 837–50
[8] Wang Z, and Ye X 2019 Space, time, and situational awareness in natural hazards: a case study of hurricane sandy with social media data *Cartography Geog. Inf. Sci.* **46** 334–46
[9] Blank G, and Lutz C 2017 Representativeness of social media in great britain: investigating Facebook, Linkedin, Twitter, Pinterest, Google+, and Instagram *Am. Behav. Sci.* **61** 741–56
[10] Hargittai E 2015 Is bigger always better? potential biases of big data derived from social network sites *Ann. Am. Acad. Pol. Soc. Sci.* **659** 63–76
[11] Clement J 2019 Facebook: number of monthly active users worldwide 2008-2019 (https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/)
[12] Crooks A, Pfoser D, Jenkins A, Croitoru A, Stefanidis A, Smith D *et al* 2015 Crowdsourcing urban form and function *Int. J. Geog. Inf. Sci.* **29** 720–41
[13] Carley K M, Malik M, Landwehr P M, Pfeffer J, and Kowalchuck M 2016 Crowd sourcing disaster management: the complex nature of twitter usage in Padang Indonesia *Saf. Sci.* **90** 48–61
[14] Liu Z, Du Y, Yi J, Liang F, Ma T, and Pei T 2019 Quantitative estimates of collective geo-tagged human activities in response to typhoon Hato using location-aware big data *Int. J. Digit. Earth* pp 1–21
[15] Lu X, and Brelsford C 2014 Network structure and community evolution on twitter: human behavior change in response to the 2011 japanese earthquake and tsunami *Sci. Rep.* **4** 6773
[16] Wang Z, and Ye X 2018 Social media analytics for natural disaster management *Int. J. Geog. Inf. Sci.* **32** 49–72
[17] Maas P, Iyer S, Gros A, Park W, McGorman L, and Nayak C 2019 Facebook disaster maps: aggregate insights for crisis response & recovery *Proc. of the 25th ACM SIGKDD Int. Conf. on Knowledge Discovery & Data* (May 2019 Valencia, Spain) pp 3173
[18] Jia S, Kim S, Nghiem S, and Kafatos M eds 2018 Near-real time population movement during an emergency A Case Study on the Mendocino Complex Fire with Disaster Maps by Facebook. *AGU Fall Meeting Abstracts* **2018** pp NH23C–0856
[19] Esri 2019 Location analytics supports nonprofit's wildfire relief efforts (https://www.esri.com/en-us/landing-page/product/2019/direct-relief-case-study-v2)
[20] Schroeder A "Data for good" and the new humanitarian future 2019 (https://www.directrelief.org/2019/01/data-for-good-and-the-new-humanitarian-future/)
[21] Center for International Earth Science Information Network (CIESIN) Columbia University 2018 Gridded Population of the World, Version 4 (GPWv4) *Population Count, Revision 11* (Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC))
[22] Harris N L, Goldman E, Gabris C, Nordling J, Minnemeyer S, Ansari S *et al* 2017 Using spatial statistics to identify emerging hot spots of forest loss *Environ. Res. Lett.* **12** 024012
[23] Hamed K H 2009 Exact distribution of the Mann-Kendall trend test statistic for persistent data *J. Hydrol.* **365** 86–94
[24] Mann H B 1945 Nonparametric tests against trend *Econometrica* **13** 245–59
[25] Kendall M G 1948 *Rank Correlation Methods* (London: Oxford University Press)
[26] Getis A, and Ord J K 1992 The analysis of spatial association by use of distance statistics *Geog. Anal.* **24** 189–206
[27] Ord J K, and Getis A 1995 Local spatial autocorrelation statistics: distributional issues and an application *Geog. Anal.* **27** 286–306
[28] Esri How emerging hot spot analysis works (https://pro.arcgis.com/en/pro-app/tool-reference/space-time-pattern-mining/learnmoreemerging.htm)
[29] Center for International Earth Science Information Network (CIESIN) Columbia University 2018 Documentation for the Gridded Population of the World, Version 4 (GPWv4) Revision 11 Data Sets (Palisades NY: NASA Socioeconomic Data and Applications Center (SEDAC)) 10.7927/H45Q4T5F
[30] Metaxa-Kakavouli D, Maas P, and Aldrich D P 2018 How Social Ties Influence Hurricane Evacuation Behavior *Proc. ACM Hum. Comput. Interact.* **2** 122
[31] Cutter S L, Ash K D, and Emrich C T 2014 The geographies of community disaster resilience *Glob. Environ. Change* **29** 65–77
[32] Tyler M C H. 2018 Woolsey, Hill fires cause Spectrum internet, cable outages in Ventura County (https://www.vcstar.com/story/news/2018/11/11/woolsey-hill-fires-cause-spectrum-outages-ventura-county/1966230002/ (Accessed 11 November 2019)
[33] Palen L, and Anderson K M 2016 Crisis informatics-new data for extraordinary times *Science* **353** 224–5
[34] Pickrell J 2019 Massive Australian blazes will 'reframe our understanding of bushfire': science news (available at: https://www.sciencemag.org/news/2019/11/massive-australian-blazes-will-reframe-our-understanding-bushfire)
[35] Buckee C O, Balsari S, Chan J, Crosas M, Dominici F, Gasser U *et al* 2020 Aggregated mobility data could help fight COVID-19 *Science* **368** 145–6